
Quentin Rossy

Simulation de cas réels de reconnaissance de locuteurs au moyen du logiciel ASPIC

JUIN 2003

Cette recherche vise à analyser l'influence des paramètres d'enregistrement, de transmission et de la langue sur les performances d'une méthode de reconnaissance de locuteur représentant l'état de l'art. Le présent rapport est divisé en quatre parties.

La partie théorique décrit le cadre et le fonctionnement du système de reconnaissance automatique utilisé. La méthodologie utilisée pour interpréter l'indice, dans un cadre bayésien, est également présentée et discutée dans cette partie.

La recherche bibliographique résume les recherches effectuées à ce jour pour évaluer l'influence des paramètres analysés dans ce travail. Le nombre d'études est plutôt faible et leurs résultats sont parfois contradictoires.

La recherche expérimentale se divise en trois parties. Premièrement, une base de données satisfaisant aux conditions théoriques nécessaires pour évaluer le système a été créée. Elle est le résultat de l'enregistrement de dix locuteurs masculins. Deuxièmement, l'influence des paramètres sur les performances du système a été analysée. Finalement, l'influence du type des enregistrements de référence utilisés pour créer les modèles qui sont comparés aux traces est étudiée.

Les résultats de la recherche sont finalement analysés dans la discussion et des recommandations sont formulées dans la conclusion.

Les principaux résultats de ce travail peuvent se résumer ainsi :

1. Lorsque la trace est enregistrée analogiquement et le modèle numériquement, les performances sont fortement diminuées.
2. Lorsque la trace est transmise par le réseau mobile et le modèle par le réseau fixe, les performances sont fortement diminuées.
3. L'utilisation d'enregistrements de référence dont les conditions techniques sont similaires à la trace doit être privilégiée.
4. Des bases de données satisfaisant à chaque condition technique ne sont pas forcément disponibles. Un type de modèle utilisable dans diverses situations doit donc être déterminé. Dans ce but, l'utilisation des enregistrements numériques transmis par le réseau fixe doit être privilégié.
5. Des modèles enregistrés analogiquement ou transmis par le réseau GSM ne devraient pas être utilisés avec des traces dont les conditions techniques ne sont pas similaires.
6. L'influence de la langue (entre le français et l'allemand) est très inférieure à l'influence de paramètres techniques.
7. L'utilisation d'un modèle dont la langue est identique à celle de la trace offre des performances plus grandes.